Reversed logical models for the study of basins of attraction

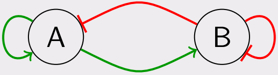
A Naldi[1], S Fueyo[2], P Monteiro[3], J Dorier[4], E Remy[5], C Chaouiya[2]

1. Univ Montpellier, now IBENS Paris
2. IGC, Lisbon
3. INESC-ID, Lisbon
4. SIB-UNIL, Lausanne
5. IML, Marseille

2017/03/13

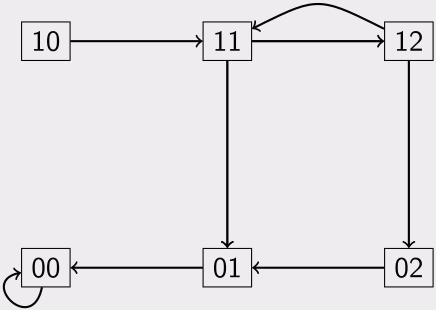
Logical Formalism

Regulatory Graph



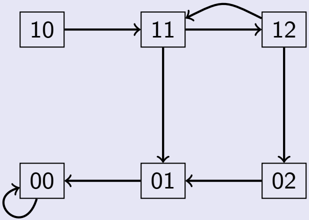
- Components
- Interactions
- Logical functions
  - $f_A = A \wedge \neg B$
  - $f_B = A \wedge \neg B : 2$

State Transition Graph

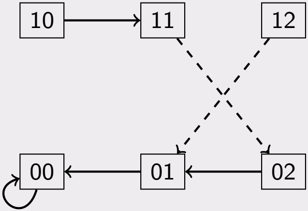


Dynamical Behaviour

Asynchronous



Synchronous



Other updatings

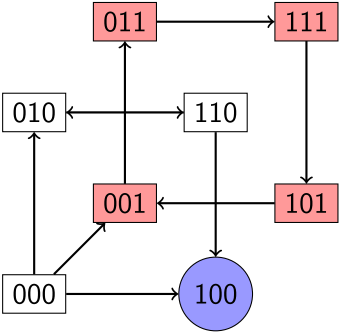
- Sequential
- Block-sequential
- Random walks

Properties

- Attractors
  stable states/oscillations
- Reachability

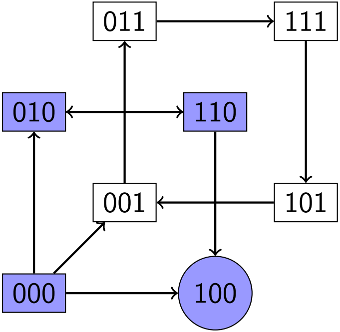
Basins of attraction

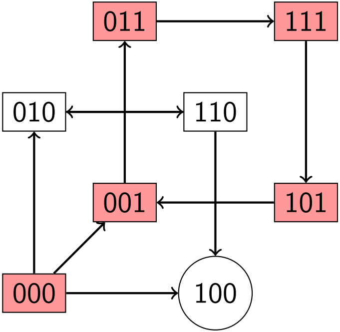
STG for a toy model with 2 attractors



Basins of attraction

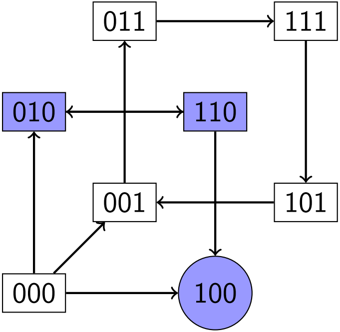
Weak basin: states from which the attractor is reachable

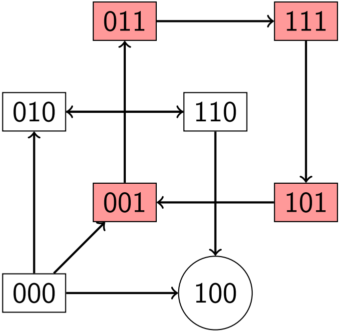




Basins of attraction

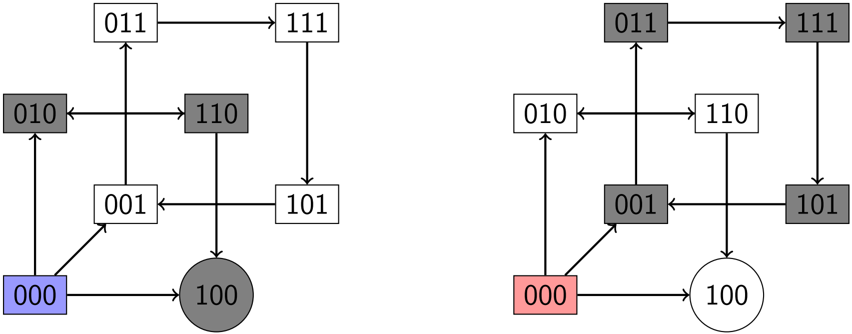
Strong basin: states from which no other attractor is reachable





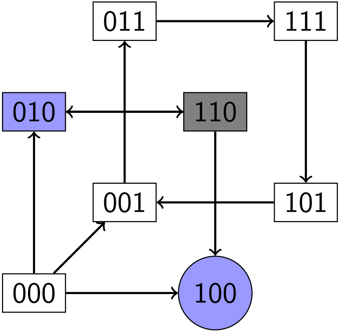
Basins of attraction

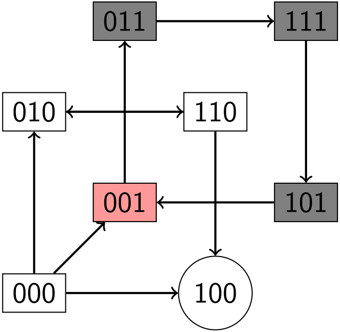
External frontier: external predecessors of the basin



Basins of attraction

Internal frontier: states of the basin with external predecessors





Identification of basins of attraction is hard

- Identification of attractors is well studied...
- ... as well as some specific reachability properties...
- ... but not on full basins of attraction.

- No formal identification methods
- Require a costly analysis of the full STG?
- Build them directly going ***backward*** from ***known attractors***

$\Longrightarrow$ ***Reversed model***: logical functions yield predecessor states

Construction of the reversed function

- x: state, i.e. vector of the levels of all components
- x_i level of the component i in state x
- $\bar{x}^i$: flip the value of x_i in state x
- $f_i(x)$: function of the component i applied to the state x
- $f_i^r(x)$: reversed function for i

Without self-regulation: $f_i(x) = f_i(\bar{x}^i)$

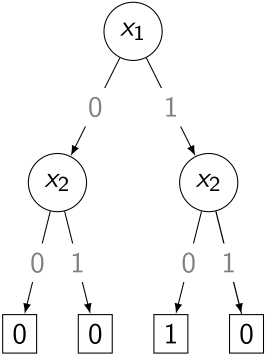
$\boxed{0} \longrightarrow \boxed{1}$ $\qquad$ $\boxed{0} \longleftarrow \boxed{1}$ $\qquad$ $f_i^r(x) = \neg f_i(x)$

With self-regulation: $f_i(x) \neq f_i(\bar{x}^i)$

$0 \qquad 1$ $\qquad$ $\boxed{0} \longleftrightarrow \boxed{1}$ $\qquad$ $f_i^r(x) = \neg f_i(\bar{x}^i)$

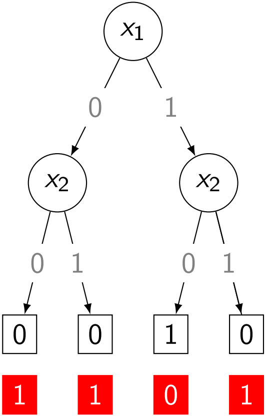
Construction of the reversed function: BDDs

$$f_1(x) = x_1 \wedge \neg x_2$$



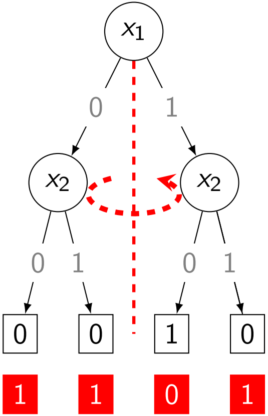
Construction of the reversed function: BDDs

$$f_1(x) = x_1 \wedge \neg x_2$$



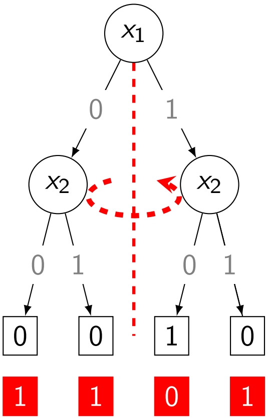
Construction of the reversed function: BDDs

$$f_1(x) = x_1 \wedge \neg x_2$$

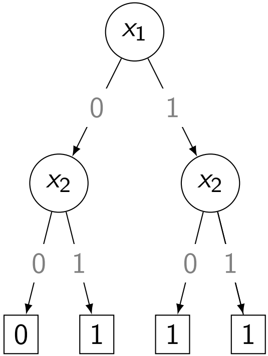


Construction of the reversed function: BDDs

$$f_1(x) = x_1 \land \neg x_2$$

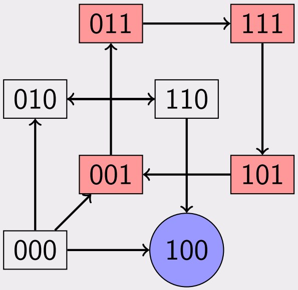


$$f_1^r(x) = x_1 \lor x_2$$

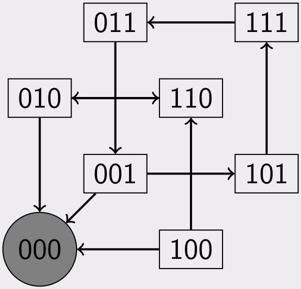


Reversed state transition graph

Original STG



Reversed STG



Properties

- Garden of eden $\Leftrightarrow$ attractors
- Reversal preserve the structure of the model and STG
- The reversed reversed model is the original model

Efficient computation of basins using boolsim

Core features of boolsim

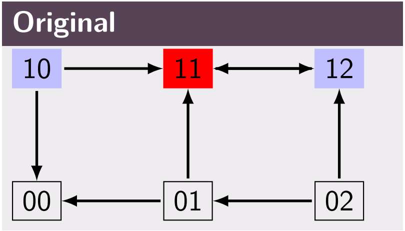
- Store a set a states as BDD
- Compute all successors of a set in one shot
- Classical combination of BDDs (AND, OR)

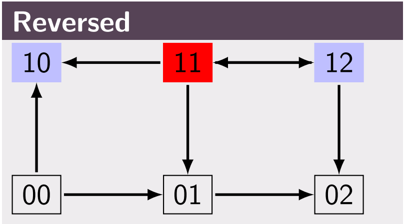
A Garg et al. *Bioinformatics* (2008)

Application to the identification of basins

- n attractors $A_1, A_2, \ldots A_n$
- W_i (weak basin of A_i): $prev \circ prev \circ \cdots \circ prev(A_i)$
- S_i (strong basin of A_i): $W_i - \{A_j \ \ \forall j \neq i\}$
- E_i (external frontier): $prev(S_i) - S_i$
- I_i (internal frontier): $next(E_i) \cap S_i$

Reversing multivalued models



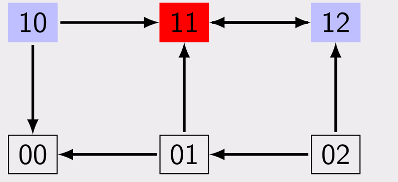


- Intermediate values are reachable from both sides
- No valid model yields the reversed STG

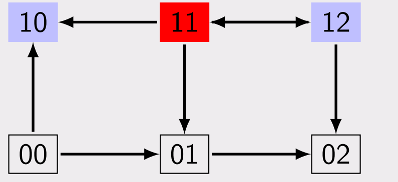
$\Rightarrow$ Most multivalued models can not be reversed

Reversing multivalued models

Original



Reversed



- Intermediate values are reachable from both sides
- No valid model yields the reversed STG

$\Rightarrow$ Most multivalued models can not be reversed

Use mapped Boolean model

- All multivalued models can be booleanized
- All Boolean models can be reversed

Booleanization

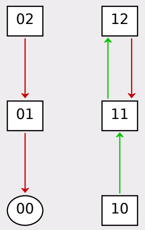
Choice of Boolean mapping

- Classical binary unsuitable: $00 \rightarrow 01 \rightarrow 10 \rightarrow 11$
- custom mapping confusing: $00 \rightarrow 01 \rightarrow 11 \rightarrow 10$
- Our pick: van Ham mapping: $000 \rightarrow 001 \rightarrow 011 \rightarrow 111$
  - Preserves the structure of the model

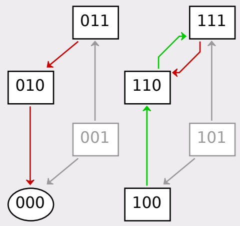
- One boolean component per activity level k
  $\Rightarrow$ Its function is f_i^k
- Introduces many non-admissible states: 100, 010, 110, 101
- Extra work to ensure "good booleanization":
  - No path from admissible to non-admissible states
  - No non-admissible states in any attractor

Booleanization and reversal

Multivalued



Booleanized

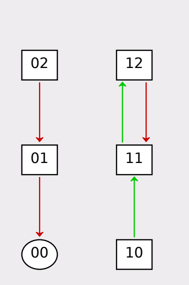


Definition of $f_i^k(x)$

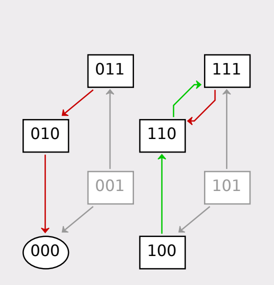
- Based on $f_i(x) \geq k$
- Restricted to escape non-admissibles

Booleanization and reversal

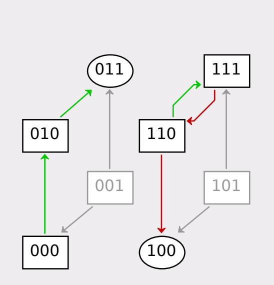
Multivalued



Booleanized



Reversed



Definition of $f_i^k(x)$

- Based on $f_i(x) \geq k$
- Restricted to escape non-admissibles
- Repeat restriction after reversal

Summary

- Construct reversed asynchronous models
  - Does NOT work for synchronous updating
  - Extend to multivalued models through boolean mapping
- Study basins of attraction
  - Requires known attractors
  - Using boolsim to compute set of reachable states
  - Identify decisive states and transitions
- Related ongoing work by Klarner & Siebert
  - Modified model checker (patch NuSMV)
- Refine reachability analysis?